AJIC special article

# Administrative coding data, compared with CDC/NHSN criteria, are poor indicators of health care–associated infections

Kurt B. Stevenson, MD, MPH,[a,b] Yosef Khan, MBBS, MPH,[a,b] Jeanne Dickman, MT, CIC,[a] Terri Gillenwater, RN, BSN,[c] Pat Kulich, RN, CIC,[a] Carol Myers, RN, BSN, CIC,[a] David Taylor, PhD,[a] Jennifer Santangelo, BA,[d] Jennifer Lundy, BS, MHA,[d] David Jarjoura, PhD,[e] Xiaobai Li, PhD,[e] Janice Shook, BS,[b] and Julie E. Mangino, MD[a,b]
Columbus, Ohio

***Background:*** ICD-9-CM coding alone has been proposed as a method of surveillance for health care-associated infections (HAIs). The accuracy of this method, however, relative to accepted infection control criteria is not known.
***Methods:*** Retrospective analysis of patients at an academic medical center in 2005 who underwent surgical procedures or who were at risk for catheter-associated bloodstream infections or ventilator-associated pneumonia was performed. Patients previously identified with HAIs by Centers for Disease Control and Prevention's National Healthcare Safety Network surveillance methods were compared with those of the same risk group identified by secondary infection ICD-9-CM codes. Discordant cases identified by only coding were all rereviewed and adjusted prior to final analysis. When coding and surveillance were both negative, a sample of patients was used to estimate the proportion of false negatives in this group.
***Results:*** The positive predictive values (PPVs) ranged from 0.14 to 0.51 with an aggregate of 0.23, even after adjustment for additional cases detected on subsequent medical record review. The negative predictive values (NPVs) ranged from 0.91 to 1.00, with an aggregate of 0.96. The estimates of the true variance of PPVs and NPVs across surgical procedures were small (0.0129, standard error, 0.009; 0.000145, standard error, 0.00019, respectively) and could be mostly explained by variation in prevalence of surgical site infections.
***Conclusion:*** Administrative coding alone appears to be a poor tool to be used as an infection control surveillance method. Its proposed use for routine HAI surveillance, public reporting of HAIs, interfacility comparisons, and nonpayment for performance should be seriously questioned. (Am J Infect Control 2008;36:155-64.)

Detection and measurement of the presence of health care-associated infections (HAIs) through surveillance is a key and essential component of established and effective health care infection control programs.[1] From 1974 through 1983, the Centers for Disease Control and Prevention (CDC) carried out the seminal Study on the Efficacy of Nosocomial Infection Control (SENIC), which demonstrated, among other findings, that surveillance is an essential element of infection control.[2] The SENIC hospitals with the lowest HAI rates had both strong surveillance and prevention control programs. Other studies have further demonstrated that surveillance promotes reduction in HAIs.[3-5] The most widely accepted methodology for defining and monitoring of HAIs was developed by the CDC through the National Nosocomial Infection Surveillance (NNIS) (now the National Healthcare Safety Network [NHSN]) system).[6,7] This standardized surveillance system using uniform and widely accepted definitions and consistent processes has been shown to be of great value for interfacility comparison by the CDC/NNIS system[8-10] and the current CDC/NHSN system.[11] The CDC/NHSN system methods for expression of infection rates and risk adjustment have become a national and international standard for categorizing and benchmarking HAI rates.

From the Department of Clinical Epidemiology,[a] Ohio State University Medical Center; Division of Infectious Diseases,[b] College of Medicine; Department of Quality and Operations,[c] Ohio State University Medical Center; OSUMC Information Warehouse,[d] Ohio State University Medical Center; and Center for Biostatistics,[e] The Ohio State University, Columbus, OH.

Address correspondence to Kurt B. Stevenson, MD, MPH, Associate Professor of Medicine, Division of Infectious Diseases, Department of Internal Medicine, The Ohio State University College of Medicine, N1147 Doan Hall, 410 West 10th Avenue, Columbus, OH 43210. E-mail: kurt.stevenson@osumc.edu.

Financial disclosures: None.

Supported by the Research Foundation of the Association for Professionals in Infection Control and Epidemiology (APIC Research Foundation).

0196-6553/$34.00



doi:10.1016/j.ajic.2008.01.004

The development of HAIs has been recognized as a major patient safety issue[12-14] prompting payors and consumers to call for public reporting of HAIs by provider organizations. Public reporting of these data has been requested as necessary for comparison of performance by health care organizations and to promote improvement and enhanced patient safety. Furthermore, many states have enacted or are considering legislation mandating such reporting. The Society for Healthcare Epidemiology of America (SHEA) and the CDC Healthcare Infection Control Practices Advisory Committee (HICPAC) have both published position papers expressing concern over the methodology for surveillance and data collection that public reporting systems may employ.[15,16]

Determining infection rates from administrative databases containing *International Classification of Diseases, Ninth Revision, Clinical Modification* (ICD-9-CM) codes has been proposed and promoted as one mechanism for case finding, data collection, and reporting. These data are readily available and would allow the application of standard definitions to be applied to administrative data sets. It would appear that the use of administrative data might provide an efficient mechanism for tracking infections. Several studies, however, have questioned the accuracy of these data for this purpose.[17-25] The CDC and its Healthcare Infection Control Practices Advisory Committee (HICPAC) have recommended that discharge diagnosis codes should not be used as the sole source of infection data for public reporting.[16] Other agencies and organizations, however, have advocated the use of such coding data for infection control surveillance purposes and for obtaining a perspective on patient safety. For example, ICD-9-CM and Diagnosis Related Groups (DRG) codes have been the foundation of the Patient Safety Indicators (PSI) promoted by the Agency for Healthcare Research and Quality.[26] Specifically, PSI 7 targets selected infections because of medical care using ICD-9-CM codes of 999.3 or 996.62 in any secondary diagnosis field. The consistency of these PSI 7 codes for surveillance for catheter-associated bloodstream infections has recently been questioned.[25]

Currently, over 30 states have some legislative activity related to the reporting of HAIs. Many states plan to utilize reporting directly through the CDC/NHSN system using CDC definitions or by the application of CDC/NHSN methodology for reporting by individual facilities directly to a state agency. Some states, however, are proposing the use of administrative claims data as their reporting mechanism.[27] More recently, the Center for Medicare and Medicaid Services (CMS), through provisions made by the Deficit Reduction Act of 2005 (Pub. L. 109-171), has enacted a policy to withhold payment, effective October 2008, for certain conditions acquired during hospitalization, including specific HAIs (catheter-associated urinary tract infection, vascular catheter-associated infection, and mediastinitis after coronary artery bypass grafting).[28,29] ICD-9-CM codes, grouped as DRGs not present on admission, would be the sole method used to identify these conditions.

Given the confidence being attributed to ICD-9-CM coding for detecting HAIs by these diverse systems of reporting and surveillance, the current retrospective study was designed to compare directly the accuracy of HAIs identified by ICD-9-CM secondary infection codes to those identified by traditional epidemiologic methods as outlined by CDC/NHSN on a large group of patients receiving care at a major academic medical center during 2005.

## METHODS

### Study location and time period

The Ohio State University Medical Center (OSUMC) is a 1145-bed tertiary care, multifacility complex in Columbus, Ohio, consisting of several adjacent buildings on the main health care campus, providing highly specialized patient care plus a community-based facility located 6 miles from the main health care complex. There were 5 infection control professionals (ICPs) who conducted infection control surveillance and implemented control interventions for all of these facilities during the study period.

OSUMC maintains a large array of clinical patient data, categorized within data marts in its information warehouse, facilitating ad hoc queries for the data retrieval required for this study. This study was a retrospective review of infection control surveillance and administratively coded data collected from January 1, 2005, to December 31, 2005 (calendar year 2005 [CY2005]), at OSUMC. This entire project was approved and monitored by the Institutional Review Board of The Ohio State University's Office of Responsible Research Practices.

### Targeted HAI surveillance

In CY2005, infection control surveillance at OSUMC consisted, in part, of monitoring for surgical site infections (SSIs) for the following targeted surgeries: coronary artery bypass grafting, peripheral vascular, colorectal, head and neck, hysterectomy, laminectomy and spinal fusion (combined as "spinal surgeries"), craniotomy, ventricular shunt placement, and total knee and hip replacements. Catheter-associated bloodstream infections (CA-BSIs) were monitored during CY2005 in several units. Likewise, ventilator-associated pneumonias (VAPs) were also monitored but in only 1 intensive care unit and for only a portion of that year. All HAIs were originally determined by routine

**Table 1.** ICD-9 CM procedure codes for classifying patients at risk for infection

| Procedure | ICD-9 procedure codes |
|---|---|
| CABG | 36.1, 36.10, 36.11, 36.12, 36.13, 36.14, 36.15, 36.16, 36.17, 36.19, 36.2 |
| Peripheral vascular | 38, 38.02, 38.03, 38.04, 38.05, 38.06, 38.07, 38.08, 38.09, 38.10, 38.12, 38.13, 38.14, 38.15, 38.16, 38.18, 38.3, 38.32, 38.33, 38.34, 38.35, 38.36, 38.37, 38.38, 38.39, 38.40, 38.42, 38.43, 38.44, 38.45, 38.46, 38.47, 38.48, 38.49, 38.7, 38.8, 38.82, 38.83, 38.84, 38.85, 38.86, 38.87, 38.88, 38.89, 39.0, 39.01, 39.02, 39.03, 39.04, 39.05, 39.06, 39.07, 39.08, 39.09, 39.10, 39.11, 39.12, 39.13, 39.14, 39.15, 39.16, 39.17, 39.18, 39.19, 39.20, 39.21, 39.22, 39.23, 39.24, 39.25, 39.26, 39.28, 39.29, 39.50, 39.51, 39.52, 39.53, 39.54, 39.55, 39.56, 39.57, 39.58, 39.59, 39.7, 39.71, 39.72, 39.73, 39.74, 39.75, 39.76, 39.77, 39.78, 39.79 |
| Colorectal | 45, 45.0, 45.00, 45.03, 45.41, 45.49, 45.5, 45.52, 45.71, 45.72, 45.73, 45.74, 45.75, 45.76, 45.79, 45.8, 45.9, 45.92, 45.93, 45.94, 45.95, 46, 46.0, 46.03, 46.04, 46.1, 46.11, 46.13, 46.14, 46.43, 46.52, 46.75, 46.76, 46.91, 46.92, 46.94, 48.5, 48.61, 48.62, 48.63, 48.64, 48.65, 48.66, 48.67, 48.68, 48.69 |
| Head and neck | 30.1, 30.4, 40.4, 40.41, 40.42 |
| Hysterectomy | 68.31, 68.39, 68.4, 68.51, 68.59, 68.6, 68.7 |
| Spinal surgeries | 03.01, 03.02, 03.09, 80.5, 80.59, 80.51, 81.01, 81.02, 81.03, 81.04, 81.05, 81.06, 81.07, 81.08, 81.03, 81.31, 81.32, 81.33, 81.34, 81.35, 81.36, 81.37, 81.38, 81.39, 81.62, 81.63, 81.64, 84.51, 84.52 |
| Craniotomy | 01.2, 01.21, 01.22, 01.23, 01.24, 01.25, 01.26, 01.27, 01.30, 01.31, 01.32, 01.39, 01.40, 01.41, 01.42, 01.50, 01.51, 01.52, 01.53, 01.59, 02.91, 02.92, 07.5, 07.51, 07.52, 07.53, 07.54, 07.59, 07.6, 07.61, 07.62, 07.63, 07.64, 07.65, 07.68, 07.69, 07.7, 07.71, 07.72, 07.79, 38.01, 38.11, 38.31, 38.41, 38.51, 38.61, 38.81, 02.11, 02.12, 02.13, 02.14 |
| Ventricular shunt | 02.2, 02.31, 02.32, 02.33, 02.34, 02.35, 02.39, 02.42, 02.43 |
| Knee and hip | 81.51, 81.52, 81.53, 81.54, 81.55 |
| Central venous catheter placement | 38.93 |
| Mechanical ventilation | 31.1, 31.2, 31.29, 31.21, 96.04, 96.7, 96.70, 96.71, 96.72 |

standard surveillance in CY2005 to have met CDC/NHSN criteria as established and in practice during 2005.[11]

## Determination of population at risk for targeted infections

For each of the targeted SSIs, the exact population at risk was recreated and determined by querying the information warehouse data marts using the ICD-9-CM procedure codes outlined in Table 1. These procedure codes exactly represented those used during CY2005 to conduct infection control surveillance and to calculate denominator data to determine infection rates. Thus, any patient on targeted units in which infection control surveillance was being conducted in CY2005 and who had one of these procedures performed was retrospectively identified. Based on the restrictions placed on using prisoner data by the OSU Institutional Review Board, all procedures performed on or devices placed in prisoners were excluded from the denominator data for the entire study; likewise, infections identified in prisoners were also excluded from the numerator data for this entire study.

The populations at risk for the development of CA-BSIs were determined by identifying those patients who had a central venous catheter placed (ICD-9 CM code 38.93) (Table 1) plus a positive blood culture obtained after catheter placement. Patients at risk for the development of VAP were identified by querying the information warehouse for patients who had any ICD-9 ventilator procedure codes (31.1, 31.2, 31.29, 31.21, 96.04, 96.7, 96.70, 96.71, 96.72) (Table 1) in the targeted intensive care unit during the same time period when infection control surveillance was conducted.

## Targeted administrative coding for infections

To avoid investigator bias, ICD-9-CM discharge codes previously selected by independent groups were reviewed and subsequently utilized as the comparison group for this study. Officials in the state of Pennsylvania, through the Pennsylvania Health Care Cost Containment Council (PHC4), have been utilizing a large group of codes as part of their public reporting for HAIs (http://www.phc4.org). In 2003, Pennsylvania legislation implemented the mandatory reporting of HAIs. In 2004, the PHC4 began the directive for all hospitals to report, through billing record format of patient discharges, hospital-wide CA-BSIs, catheter-associated urinary tract infections, VAP, and SSIs (cardiovascular, neurosurgical, and orthopedic procedures).[19,20] The PHC4 defined more than 140 secondary ICD-9 CM infection codes to identify HAIs as outlined in Table 2. After a review of ICD-9-CM codes utilized by states for public reporting, the PHC4 codes were thought to represent the most comprehensive selection. Thus, these sets of secondary infection codes for SSIs, CA-BSIs, and VAPs were directly employed to identify potential cases of HAIs in this study by administrative coding.

## Determination of patients with an infection as identified by coding data

A query of the OSUMC information warehouse databases was performed using all of the SSI administrative

**Table 2.** ICD-9 CM discharge codes used as screening for infection*

| Type of infection | ICD-9 infection screening codes |
|---|---|
| Surgical site infections | 994.2, 996.61, 996.62, 996.63, 996.66, 996.67, 996.71, 996.72, 998.0, 998.31, 998.32, 998.51, 998.59, 998.6, 999.83, 999.3, 320.81, 320.82, 320.89, 320.0, 320.1, 320.2, 320.3, 320.7, 320.9, 321.0, 321.1, 321.2, 321.3, 321.4, 321.8, 322.0, 322.1, 322.2, 322.9, 324.0, 324.1, 324.9, 420.90, 420.91, 420.99, 421.9, 422.90, 422.91, 513.1, 519.2, 682.1, 682.2, 682.3, 682.4, 682.6, 682.7, 682.9, 728.0, 730.00, 730.01, 730.02, 730.03, 730.04, 730.05, 730.06, 730.07, 730.08, 730.09, 730.20, 730.21, 730.22, 730.23, 730.24, 730.25, 730.26, 730.27, 730.28, 730.29, 730.30, 730.31, 730.32, 730.33, 730.34, 730.35, 730.36, 730.37, 730.38, 730.39, 730.90, 730.91, 730.92, 730.93, 730.94, 730.95, 730.96, 730.97, 730.98, 730.99, 890.0, 890.1, 890.2, 891.0, 891.0, 891.1, 891.2, 894.0, 894.1, 894.2 |
| Bloodstream infections | 0.38.0, 038.10, 038.11, 038.19, 038.2, 038.3, 038.40, 038.41, 038.42, 038.43, 038.44, 038.49, 038.8, 038.9, 790.7, 995.90, 995.91, 995.92, 480.0, 480.1, 480.2, 480.3, 480.8, 480.9, 481, 482.0, 482.1, 482.2, 482.30, 482.31, 482.32, 482.39, 482.40, 482.41, 482.82, 482.83, 482.84, 482.89, 482.9, 483.0, 483.1, 483.8, 485, 486, 487.0, 482.49, 482.81 |
| Pneumonia | 480.0–480.3, 480.8, 480.9, 481, 482.0–482.2, 482.30–482.32, 480.39–482.41, 482.82–482.84, 482.89, 482.9, 483.0, 483.1, 483.8, 485, 486, 487.0, 482.49, 482.81 |

*Adapted from Pennsylvania Health Care cost Containment Council (PHC4).

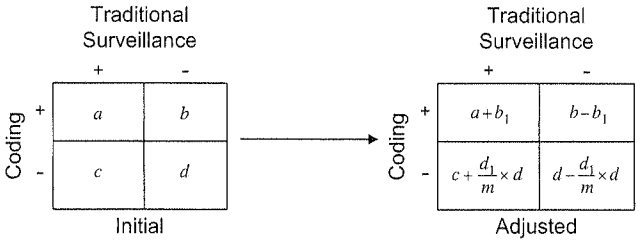


**Fig 1.** Schematic representation of infection case classifications from both coding and traditional surveillance methods for initial and adjusted analyses. Identical sets of patients were queried for infections using ICD-9-CM secondary infection codes or CDC/NHSN criteria (traditional surveillance). Discordant cases (count *b*) were reviewed and, if found to meet CDC/NHSN criteria (count $b_1$), were reclassified. False negatives where neither method identified a case (count *d*) were determined by a review of random cases and reclassified as true cases as illustrated. Using traditional surveillance as the reference standard, the accuracy of coding data was determined from the adjusted 2 × 2 table. See text for complete details.

codes (Table 2) for the time period January 1, 2005, to December 31, 2006, to include queries for up to 1 year after certain surgical procedures were performed. This query defined all patients coded with any of the targeted administrative ICD-9-CM infection screening codes during the study time period. Each specific group of patients undergoing a targeted surgical procedure as outlined above (Table 1) was matched against this group of patients to determine which of those having the targeted procedure had also been coded with any 1 or more ICD-9-CM codes for a suspected infection from Table 2 during the designated study time period. Patients were matched only if they had the infection code added after the date of the targeted surgery. Because of the placement of an implant or artificial device, patients with coronary artery bypass grafting, peripheral vascular, spine, craniotomy, ventricular shunt, and knee and hip joint replacement surgeries were screened for the presence of the screening infection ICD-9-CM codes for up to 1 year after the procedure, consistent with CDC/NHSN guidelines. Patients with colorectal, head and neck, and hysterectomy procedures were queried for only 30 days after the targeted procedure, consistent with CDC/NHSN guidelines. Upon completion of this step, all patients undergoing a targeted surgical procedure in CY2005 who were also categorized as having an infection based on ICD-9-CM coding from the PHC4 list of screening codes were thus identified.

Patients within the targeted units with the presence of a central venous catheter plus a positive blood culture were matched with those having 1 or more ICD-9-CM screening codes for bloodstream infection (Table 2). Patients within the targeted unit with a procedure code related to mechanical ventilation were matched with those having 1 or more ICD-9-CM screening codes for pneumonia (Table 2).

## Matching of patients identified with infection from traditional surveillance to those identified with an infection based on coding data

Utilizing a line list of all CY2005 patients previously identified with a HAI based on CDC/NHSN definitions and methodology (hereafter referred to as "traditional surveillance"), cases were systematically matched within their specific groups obtained from coding data. This matching created groups of patients for each surgical procedure type or targeted device infection: (1) cases in which both traditional surveillance and coding data agreed that there was an infection (count *a*, Fig 1), (2) cases in which coding data identified an infection and traditional surveillance did not (count *b*, Fig 1), (3) cases in which traditional surveillance identified an infection and coding did not (count

*c*, Fig 1), and (4) cases in which neither method identified an infection (count *d*, Fig 1). This step encompassed our "initial" analysis of the comparison of coding data to traditional surveillance.

### Medical record review of discordant cases and cases not detected by either surveillance method

Discordant cases identified by ICD-9-CM coding but not by traditional surveillance (count *b*, Fig 1) were each individually reviewed by ICPs to determine the presence of an infection based on CDC/NHSN definitions and methods utilized in CY2005. For all cases, this consisted of a review of *all* discordant cases except potential CA-BSI cases. Furthermore, a random sample of 40 cases (referred to as count *m*, Fig 1) from each group in which neither method detected an infection (count *d*, Fig 1) was reviewed to estimate the number of cases that should have been included in count *c*. This random sample represented approximately 10% of total cases for each group.

All medical record reviews were completed independently of any input or discussion of cases by the investigators or other members of the research team so as to not introduce any bias in the final decision regarding the presence or absence of an infection. The ICPs were instructed to follow their standard procedures for reviewing these cases and to identify as cases only those that were in compliance with CDC/NHSN definitions and methodology based on their usual practices. Upon completion of the medical record review, any of the discordant cases found to be an infection (count $b_1$ and proportion $d_1/m$, Fig 1) were reclassified for the adjusted analysis as outlined in the Data and Statistical Analysis section below. Because CA-BSI cases were also randomly sampled, the cases were also converted to a proportion for reclassification.

### Data and statistical analysis: Estimates of positive predictive values and negative predictive values of the coded surveillance method after verification of negatives through subsampling

The rationale and strategy for final data analysis are outlined. Two-by-two tables were completed in which the coded data were crossed with traditional epidemiologic surveillance (Fig 1). We assumed for the final analysis that, if traditional surveillance determined the presence of an infection, according to CDC/NHSN criteria, then it was truly positive. We did not assume that traditional surveillance provides gold standard classification without verifying negatives. Figure 1 refers to these as counts *a* and *c*. When traditional surveillance was not positive (counts *b* and *d*), the number of false negatives was verified among these counts and added to counts *a* and *c*. Count *b* was small enough to determine completely for all patients the number that was missed by traditional surveillance, with the exception of CA-BSI cases from which a random sample of 50 cases was obtained. We refer to this number in the second table of Fig 1 as $b_1$. As one would expect across all procedures, there were many patients who were negative on coding and negative on traditional surveillance (count *d*). From those patients, we took samples of 40 patients of the *d*, which we refer to as *m*, and used them to estimate the proportion of that cell that were false negatives ($d_1/m$). Figure 1 shows how the initial stage table was adjusted to incorporate these corrections for false negatives. From the data in the final adjusted $2 \times 2$ tables, the calculations of positive predictive value (PPV) and negative predictive value (NPV) with coding data classified as the "test" compared to adjusted traditional surveillance as the "reference standard" were calculated according to standard epidemiologic methods.[30]

Sampling from *d* to estimate the false-negative proportion makes standard estimates of variance of summary statistics like PPV and NPV inapplicable. We developed a model in which we assumed that sampling was fixed within each of the negative and positive categories of the electronic coding results, which is a standard assumption in such tables. In other words, we assumed $a+b$ fixed and $c+d$ fixed. To account for the subsampling within the negative cells, we assumed the sample size was fixed to *m* for the coding negatives and that *d* would always be larger than *m* (Li X, et al, manuscript in preparation). We obtained maximum likelihood estimates (MLE) of PPV and NPV and their variances. For comparison purpose, we also provide the variance estimates from the standard approach. Note that we also allowed for sampling from count *b*, even though it was not needed in this study except for CA-BSI. Thus, the MLEs allow for subsampling for verification from *b* as well as *d*.

To evaluate the variability of PPVs and NPVs across all surgical procedures, we estimated the true variance of these not explainable by sampling error. The true PPVs could be all the same across procedures, and they vary only because of sampling error, or they vary systematically across procedures beyond what can be explained by sampling error. The CA-BSIs and VAPs were not included in this systemic analysis because they could not be directly compared with the surgical procedures or to each other because of the unique nature of these device-related infections.

## RESULTS

During CY2005, there were a total of 3882 surgical procedures performed at OSUMC for which traditional

**Table 3.** Measures of accuracy of coding data compared with traditional surveillance after adjustment

| | PPV | SDL | SDS | NPV | SDL | SDS | Prevalences |
|---|---|---|---|---|---|---|---|
| Infection type | | | | | | | |
| Surgical site infections | | | | | | | |
| Coronary bypass grafting | 0.42 | 0.067 | 0.067 | 0.96 | 0.025 | 0.009 | 0.08 |
| Peripheral vascular surgery | 0.35 | 0.050 | 0.050 | 0.97 | 0.009 | 0.009 | 0.09 |
| Colorectal | 0.51 | 0.058 | 0.058 | 0.95 | 0.025 | 0.01 | 0.11 |
| Head and neck | 0.19 | 0.069 | 0.069 | 0.99 | 0.008 | 0.008 | 0.03 |
| Hysterectomy | 0.46 | 0.096 | 0.094 | 0.98 | 0.008 | 0.008 | 0.05 |
| Spinal surgeries | 0.14 | 0.050 | 0.050 | 1.00 | 0.001 | 0.001 | 0.01 |
| Craniotomy | 0.22 | 0.080 | 0.080 | 1.00 | 0.005 | 0.005 | 0.03 |
| Ventricular shunt | 0.29 | 0.099 | 0.099 | 0.99 | 0.008 | 0.008 | 0.06 |
| Knee and hip replacements | 0.22 | 0.047 | 0.047 | 0.99 | 0.004 | 0.004 | 0.03 |
| Others | | | | | | | |
| Catheter-associated BSI | 0.15 | 0.015 | 0.015 | 0.91 | 0.008 | 0.009 | 0.11 |
| Ventilator–associated pneumonia | 0.34 | 0.075 | 0.074 | 0.91 | 0.023 | 0.023 | 0.15 |
| Totals | 0.23 | | | 0.96 | | | |

*PPV*, positive predicitive value; *NPV*, negative predicitive value; *SDL*, standard error for the estimated PPV/NPV from the model approach; *SDS*, standard error for the estimated PPV/NPV from the traditional approach.

infection control surveillance was conducted using standard CDC/NHSN methods. Additionally, there were an estimated 1599 patients at risk for CA-BSIs and 193 patients at risk for VAP in the units at OSUMC in which traditional surveillance was conducted based on the denominator definitions applied. Among the surgical procedures, 457 (12%) were initially identified as having a SSI by coding data and 144 (4%) by traditional infection control surveillance methods. Among patients with a central venous catheter placed and with a positive blood culture, 569 (36%) were initially identified as having a CA-BSI by coding data and 150 (9%) by traditional infection control surveillance. Among patients with the presence of mechanical ventilation and intubation, 41 (21%) were initially identified as having a VAP by coding data and 24 (12%) by traditional infection control surveillance. All of these data constituted the "initial" data for analysis of coding accuracy (Fig 1).

Upon completion of the initial data collection, there were 363 discordant cases (count *b*, Fig 1) among surgical procedures for which subsequent medical record review was performed. These were cases that coding identified as infections but were not identified by initial infection control surveillance. Among these, 55 of 363 (15%) were subsequently found to meet CDC/NHSN criteria for an infection but were missed by initial surveillance (count $b_1$, Fig 1). There were 485 discordant cases (count *b*, Fig 1) among potential CA-BSI cases, and a random sample of 50 of these was reviewed. Among these, none were found to have a CA-BSI as determined by CDC/NHSN criteria (count $b_1$, Fig 1). There were 31 discordant cases among VAP cases (count *b*, Fig 1), and these were all reviewed. There were 4 (13%) cases found to have a VAP according to CDC/NHSN definitions (count $b_1$, Fig 1). Overall, there were 879 discordant cases that were identified as infections by coding but not traditional infection control surveillance. Combining the review of all surgical and pneumonia cases and extrapolating the random sample of BSI cases to the total, only 59 of 879 (7%) of these cases were missed by initial surveillance and determined to have actually been infections based on CDC/NHSN definitions. These cases were reclassified for the final adjusted analysis ($a + b_1$, $b$- $b_1$, Fig 1).

There were 3375 surgical cases, 964 patients with central venous catheter and positive blood cultures, and 138 ventilated patients in whom neither surveillance method identified an infection (count *d*, Fig 1). To assess whether any cases were missed among these patients, a random sample of 40 cases of each surgical procedure and device category were reviewed. No additional infections were found in all groups with the exception of 1 case each for coronary artery bypass grafting and colorectal surgery patients and 1 case among patients with central venous catheter and positive blood culture. For each of these categories, it was estimated that 1 of 40 (2.5%) ($d_1/m$, Fig 1) cases should have been classified as infections by traditional infection control surveillance. These cases were adjusted based on the total ($c + d_1/m \times d$, Fig 1) and were reclassified for the final adjusted analysis.

The accuracy of coding data compared with traditional infection control surveillance, calculated as PPV and NPV, based on these adjusted data are outlined in Table 3. Generally, the PPVs are low and the NPVs are very high and near 1.0 across all of the infections; PPVs ranged from 0.14 to 0.51 and NPVs from 0.91 to 1.00. The PPV was highest for colorectal surgeries and lowest for spinal surgeries. Overall, the aggregate PPV of coding data was only 0.23 for all cases when compared with traditional infection control surveillance for

accuracy, even after adjustment for cases that were initially missed by traditional surveillance as determined by subsequent medical record review. The standard error estimates from the MLE approach (SDL, Table 3) and the standard approach (SDS, Table 3) agreed well for PPV given that all of these cases were all verified by medical record review. However, there are marked differences between the 2 standard error estimates (SDL and SDS, Table 3) for NPV because of the verification subsampling within this group (count *d*, Fig 1), which adds sampling error to the estimate of NPV.

A comparison to determine statistical differences of PPV and NPV across surgical procedures was conducted. For PPV, the estimate of the true variance component for the surgical procedure random effect was 0.0129 with a standard error of 0.009. Thus, there is no strong evidence to suggest that the PPVs vary systematically across types of surgical procedures. Variation in the PPVs is a function of variation in infection prevalence. The logit of a PPV breaks it up into the sum of a logit of prevalence and the log of the likelihood ratio (log [sensitivity/[1-specificity]). The variation in PPVs can, therefore, be broken into that because of prevalence and that because of accuracy independent of prevalence. The estimate of the true variance of the logit of PPV is totally explained by prevalence differences across surgical procedures. For variation in NPVs, we also failed to find clear evidence of true variance among surgical procedures. This is not surprising because they are all between 0.95 and 1.00. The estimate of the true variance of NPVs was 0.000145 with a standard error of .00019. Applying the logit transformation, approximately two thirds of this small variance could be explained by variation in prevalence.

After the medical record review was complete, ICPs indicated that a common reason that cases may have been initially missed by standard surveillance but identified on subsequent medical record review was that these additional cases were identified based on clinical findings and not on culture results. If microbiology cultures are a primary case finding method, then some of these cases would have been missed. A positive microbiologic culture is commonly used as an initial screen for potential HAIs, especially SSIs. Despite this case-finding disparity, only a small number of discordant cases rereviewed actually met the CDC/NHSN criteria for true infection (59/879, 7%). This indicates that initial traditional surveillance generally performed well by not missing a significant number of infection cases.

After medical record review, ICPs attempted to ascertain reasons that cases were identified by coding data but not confirmed by traditional infection control surveillance methods, and these are outlined in Table 4. These categories are not mutually exclusive, and more than 1 reason may be present for some cases. In many cases (34%), no infection was identified on subsequent medical record review; in an equivalent number (32%), an infection was identified but at a site different from the site of the targeted infection. In several cases (15%), an infection was detected but was found to be community acquired and not hospital acquired. In a small number of cases (3%), an infection may have been detected, but it was for a procedure performed at another facility and would not have been attributed to our facility based on standard infection control surveillance practices. In a number of cases (21%), other varied reasons were cited by the ICPs reviewing the cases. Based on CDC/NHSN methods, none of these reasons, therefore, would have met criteria for classification as an HAI at our facility.

**Table 4.** Specified reasons that coding misclassified infections based on medical record review

| | |
|---|---|
| No infection identified, n | 132 (34%) |
| Infection identified but at a site different from the targeted site, n | 123 (32%) |
| Procedure performed at a different facility, n | 11 (3%) |
| Community-acquired, not healthcare-associated infection, n | 59 (15%) |
| Other, n | 77 (21%) |

Note. n = 384.

## DISCUSSION

This study was designed to provide the most direct and equitable comparison between ICD-9-CM coding data obtained from administrative databases and traditional infection control surveillance methods as practiced by most hospitals in the United States. The identical set of patients for which infection control surveillance had been previously conducted was used for comparison. Patients identified by coding data, but not CDC/NHSN criteria, as having HAIs were all reviewed and reclassified if missed initially. Only a small number of these discordant cases (59/879, 7%) were subsequently found to meet CDC/NHSN criteria, suggesting that the case-finding methods outlined by the CDC/NHSN are reasonably robust. Based on the summary results from this study, classification of HAIs by ICD-9-CM secondary infection codes, when compared with classification of infections by standard infection surveillance methods using CDC/NHSN definitions and methods, is very imprecise with an aggregate PPV of only 0.23. Thus, 3 out of 4 HAIs as detected by coding data, on average, would not meet standard CDC/NHSN definitions and criteria.

Given that coding data has been proposed as a surveillance tool for future public reporting and nonpayment, the PPV as a measure of the accuracy of coding data appears to be the most critical measure and is

the one employed in this study. The PPV varied by surgical procedure (0.14 to 0.51), but additional analysis confirmed that this variation across surgical procedures was not significant. Rather than sampling all intensive care unit patients with codes for bloodstream infections, this broad group of patients was further narrowed to those with central venous catheters and positive blood cultures. Despite these efforts to give coding data the best opportunity for identification of HAIs, the PPV value for CA-BSIs remained quite low at 0.15. Thus, ICD-9-CM discharge codes for all HAIs consistently overreported the number of infections in contrast to CDC/NHSN criteria. The NPV of the ICD-9-CM codes compared with traditional surveillance were high, suggesting that ICD-9-CM codes could be considered for screening of patient data sets as a case-finding method, but, as recently outlined by a CDC and HICPAC statement on the public reporting of HAIs,[16] they should not used as the sole method of surveillance. The utility of coding data for screening of HAIs may be further limited, however, because of the inherent time delay in the availability of such data, often not until many weeks to months after discharge.

An additional interesting feature of surveillance with ICD-9-CM discharge codes was noted. The ICD-9-CM codes for SSIs were selected by the PHC4 to be specific for cardiovascular, neurosurgical, and orthopedic SSIs. They were, however, applied in our study against a more diverse group of surgical procedures. These codes, surprisingly, had the highest PPV for procedures (colorectal, hysterectomy) not previously included in the PHC4 reporting. Although the differences in PPV across surgical procedures could be explained away by SSI prevalence, ICD-9-CM coding may lack selectivity for surveillance of specific infection types or targeted procedures.

Previous studies have suggested similar concerns about the accuracy of discharge coding data for identifying and classifying HAIs. For example, Moro and Morsillo examined a large number of patients from a regional database of postoperative infections.[17] Infections were identified prospectively using CDC/NHSN definitions, including discharge surveillance with telephone interviews and mailings. These data were compared with the ICD-9-CM discharge codes assigned to these patients. The sensitivity of coding was 10% when codes for a postoperative infection were used. This only increased to 19% and 21%, respectively, when broader sets of ICD-9-CM codes were applied.

Another study examined different combinations of ICD-9-CM codes to identify patients with pneumonia.[18] A rigorous reference standard was developed to classify patients with pneumonia based on symptoms, chest radiographs, computerized decision support system, and use of the term *pneumonia* in the admission or discharge reports. The sensitivity of claims data ranged from 47.8% to 66.2% with PPVs of 72.6% to 80.8%. Despite these higher PPVs, these authors still concluded that their ICD-9-CM coding algorithms were imprecise.[18] Two previous studies examining hospitalized patients with pneumonia showed similar results.[31,32]

Two recent studies from Pennsylvania hospitals detail the experience with reporting infections based on the PHC4 methodology using the same ICD-9-CM secondary infection codes employed in this study.[19,20] The first examined a 9-month time period in 2004 at Children's Hospital of Philadelphia.[19] A retrospective cross-sectional review of cases was performed comparing hospital infection control surveillance data with patients identified by ICD-9-CM codes specified by the PHC4. A random sample of discordant cases was reviewed. The sensitivity of cases found by administrative data compared with infection control data was 61% with a PPV of 20%.

The second study was reported from the Penn State Milton S. Hershey Medical Center at which one fourth of CY2004 data was reviewed.[20] Cases identified by PHC4 ICD-9 codes were identified, and a sample of these cases was reviewed by trained ICPs. On review, only 15% of urinary tract infection, 15% of SSIs, and 16% of VAP cases met CDC/NHSN definitions. Thus, approximately 85% of cases identified by PHC4 ICD-9 codes in this hospital as infections failed to meet standard CDC/NHSN definitions. These 2 studies from Pennsylvania support the conclusions from our present study.

Other investigators have examined a combination of ICD-9-CM codes with other health care data to increase their utility in the surveillance of HAIs. For example, ICD-9-CM codes combined with antimicrobial use data appear to be a more sensitive method for detecting postoperative infections.[33,34] Thus, there may be a role of ICD-9-CM codes as a part of surveillance algorithms combined with other data such as antimicrobial use data, microbiologic data, vital signs, or other clinical or laboratory data. The role of ICD-9-CM coding as an adjunct to current infection control surveillance methodology, however, is not yet fully determined.

There are some limitations to this study. All of the data are abstracted from the experience of only 1 large academic center. The surveillance practices of ICPs at our medical center may not be representative of those at other centers; however, the ICPs have all been trained in recognized and accepted APIC courses and have nearly 100 years of cumulative surveillance experience. Furthermore, OSUMC has participated in NNIS/NHSN for many years, giving these ICPs extensive experience in surveillance with CDC/NHSN methodology. All of these reasons make this limitation less likely to be significant. Additionally, the coding practices of medical record abstractors may be unique to our center, and coding accuracy may be different at other

medical centers. Nonetheless, the comparison of our results with other similar studies, especially those in the Pennsylvania hospitals using the same ICD-9-CM codes, would suggest otherwise. We examined the complete sample of patients at risk for infection in each category, reducing the chance for verification bias. Verification bias would occur if all subjects in the original population do not have an equal chance of disease verification.[35,36] Furthermore, the ICPs were completely independent in their reviews without any discussion of the specifics of the cases, reducing the chance that classification bias could be introduced.

CDC/NHSN definitions and criteria were utilized as the gold reference standard for testing the accuracy of coding data. Critics of the CDC methodology are concerned that the application of this system may systematically miss HAIs. A pilot study published in 1998 examined the accuracy of the CDC/NNIS methodology.[37] It was conducted in 2 phases to review the charts of selected intensive care unit patients who had HAIs reported to the NNIS System. In phase 1, retrospectively detected infections that matched the previously reported infections were deemed to be true infections. In phase II, 2 CDC epidemiologists reexamined a sample of charts for which a discrepancy existed. Each sampled infection either was confirmed or disallowed by the epidemiologists. The PPV for reported bloodstream infections, pneumonia, SSI, urinary tract infection, and other sites was 87%, 89%, 72%, 92%, and 80%, respectively; the sensitivity was 85%, 68%, 67%, 59%, and 30%, respectively; and the specificity was 98.3%, 97.8%, 97.7%, 98.7%, and 98.6%, respectively. The authors concluded that, when NNIS hospitals in the study reported a HAI, the infection most likely was a true infection and that they infrequently reported conditions that were not infections.

This study has significant implications for public reporting and the nonpayment for performance. Based on our results, coding data will overreport HAIs, on aggregate, by 3-to 4-fold. This concern is illustrated by reports from the PHC4 that have compared HAIs reported directly by Pennsylvania hospitals to those detected by coding data. Examination of data for 2004 indicated 11,668 HAIs identified by traditional surveillance methods compared with 115,000 HAIs documented from coding data.[38] Analysis of coded billing data revealed 10- to 20-fold more HAI episodes than were recorded during surveillance. Much of this discrepancy could be explained by the reasons identified by our ICPs as detailed in Table 4: infections identified but at a site different from the targeted site for coding, the procedure was performed at a different facility than the one reporting, or the infection was acquired in the community (Table 4). In many cases, no infection was actually identified. This is likely explained by different methods and criteria applied by coders where data documented by the physician are only utilized. A previous assessment of the PHC4 analysis independently identified many of these same explanations for the difference in outcomes between coding and traditional surveillance data.[39]

If HAI rates calculated from ICD-9-CM codes overreport cases by as much as 75%, then reductions noted in coding rates attributed to performance improvement interventions may merely be due to variations in coding and not represent any true reduction in infection rates. A recent systematic review suggested that public reporting is ineffective to improve health care performance.[40] Thus, inaccuracies in rates for the sake of public reporting may not significantly impact consumer choice or hospital performance. In the context of nonpayment for performance, however, these discrepancies can have significant implications. Based on our data, 3 of 4 cases may have payment withheld or reduced erroneously by Medicare. Additionally, cost data calculated using ICD-9-CM codes, and not CDC/NHSN criteria, may significantly inflate the economic impact of HAIs both in estimated expenses and projected cost savings from interventions. This study emphasizes the need for an accurate measure of HAI rates to assess accurately their true impact within the health care system.

In summary, this comprehensive study examined the ability of a large group of ICD-9-CM discharge codes to identify correctly the infections occurring after 9 different types of surgical procedures and for 2 common HAIs (CA-BSI and VAP). The PPVs were very low in all cases, suggesting that ICD-9-CM codes performed poorly and were imprecise. Thus, administrative coding alone appears to be a poor tool for widespread use as a surveillance method, especially for interfacility comparisons, public reporting, and nonpayment for performance. Accurate measures of HAIs are needed for these purposes.

## References


1. Gaynes RP. Surveillance of nosocomial infections: a fundamental ingredient for quality. Infect Control Hosp Epidemiol 1997;18:475-8.
2. Haley RW, Culver DH, White JW. The efficacy of infection surveillance and control programs in preventing nosocomial infections in US hospitals. Am J Epidemiol 1985;121:182-205.
3. Centers for Disease Control and Prevention. Public health focus: surveillance, prevention, and control of nosocomial infections. MMWR 1992;41:783-7.
4. Roy MC, Perl TM. Basics of surgical-site infection surveillance. Infect Control Hosp Epidemiol 1997;18:659-68.
5. Misset B, Timsit JF, Dumay MF, Garrouste M, Chalfine A, Flouriot I, et al. A continuous quality-improvement program reduces nosocomial infection rates in the ICU. Intensive Care Med 2004;30:395-400.
6. Gaynes RP, Culver DH, Emori TG, Horan TC, Banerjee SN, Edwards JR, et al. The National Nosocomial Infections Surveillance System: plans for the 1990s and beyond. Am J Med 1991;91(Suppl 3B):S116-20.

7. Centers for Disease Control and Prevention. Monitoring hospital-acquired infections to promote patient safety-United States, 1990-1999. MMWR 2000;49:149-53.
8. Archibald LK, Gaynes RP. Hospital-acquired infections in the United States: the importance of interhospital comparisons. Infect Dis Clin North Am 1997;11:245-55.
9. Garner JS, Jarvis WR, Emori TG, Horan TC, Hughes JM. CDC definitions for nosocomial infections, 1988. Am J Infect Control 1988;16: 28-40.
10. Horan TC, Gaynes RP, Martone WJ, Jarvis WR, Emori TG. CDC definitions of nosocomial surgical site infections, 1992: a modification of CDC definitions of surgical wound infections. Infect Control Hosp Epidemiol 1992;13:606-8.
11. Horan T, Gaynes R. Surveillance of nosocomial infections. In: Hospital epidemiology and infection control. Philadelphia: Williams & Wilkins; 2004. p. 1659-702.
12. Burke JP. Infection control—a problem for patient safety. N Engl J Med 2003;348:651-6.
13. Gerberding JL. Hospital-onset infections: a patient safety issue. Ann Intern Med 2002;137:665-70.
14. Pittet D, Donaldson L. Challenging the world: patient safety and health care-associated infection. Int J Qual Health Care 2006;18:4-8.
15. Wong ES, Rupp ME, Mermel L, Perl TM, Bradley S, Ramsey KM, et al. Public disclosure of healthcare-associated infections: the role of the Society for Healthcare Epidemiology of America. Infect Control Hosp Epidemiol 2005;26:210-2.
16. McKibben L, Horan TC, Tokars JI, Fowler G, Cardo DM, Pearson ML, et al. Guidance on public reporting of healthcare-associated infections: recommendations of the Healthcare Infection Control Practices Advisory Committee. Infect Control Hosp Epidemiol 2005;26:580-7.
17. Moro ML, Morsillo F. Can hospital discharge diagnoses be used for surveillance of surgical-site infections? J Hosp Infect 2004;56:239-41.
18. Aronsky D, Haug PJ, Lagor C, Dean NC. Accuracy of administrative data for identifying patients with pneumonia. Am J Med Qual 2005; 20:319-28.
19. Sherman ER, Heydon KH, St. John KH, Teszner E, Rettig SL, Alexander SK, et al. Administrative data fail to accurately identify cases of healthcare-associated infection. Infect Control Hosp Epidemiol 2006;27: 332-7.
20. Julian KG, Brumbach AM, Chicora MK, Houlihan C, Riddle AM, Umberger T, et al. First year of mandatory reporting of healthcare-associated infections, Pennsylvania: an infection control-chart abstractor collaboration. Infect Control Hosp Epidemiol 2006;27:926-30.
21. Romano PS, Chan BK, Schembri ME, Rainwater JA. Can administrative data be used to compare postoperative complication rates across hospitals? Med Care 2002;40:856-67.
22. Wright SB, Huskins WC, Dokholyan RS, Goldmann DA, Platt R. Administrative databases provide inaccurate data for surveillance of long-term central venous catheter-associated infections. Infect Control Hosp Epidemiol 2003;24:946-9.
23. Sands KE, Yokoe DS, Hooper DC, Tully JL, Horan TC, Gaynes RP, et al. Detection of postoperative surgical-site infections: comparison of health plan-based surveillance with hospital-based programs. Infect Control Hosp Epidemiol 2003;24:741-3.
24. Madsen KM, Schonheyder HC, Kristensen B, Nielsen GL, Sorensen HT. Can hospital discharge diagnosis be used for surveillance of bacteremia? A data quality study of a Danish hospital discharge registry. Infect Control Hosp Epidemiol 1998;19:175-80.
25. Stone PW, Horan TC, Shih HC, Mooney-Kane C, Larson E. Comparisons of health care-associated infections identification using two mechanisms for public reporting. Am J Infect Control 2007;35:145-9.
26. Agency for Healthcare Research and Quality. Patient Safety Indicators. Technical Specifications. AHRQ Quality Indicators. Available at: http://www.qualityindicators.ahrqgov/downloads/psi/psi_technical_specs_v31.pdf. March 2007. Accessed November 7, 2007.
27. Edmond MB, Bearman GM. Mandatory public reporting in the USA: an example to follow? J Hosp Infect 2007;65(Suppl 2):182-8.
28. Medicare program: changes to the hospital inpatient prospective payment systems and fiscal year 2008 rates. Fed Regist 2007;72: 47379-428.
29. Rosenthal MB. Nonpayment for performance? Medicare's new reimbursement rule. N Engl J Med 2007;357:1573-5.
30. Weiss NS. Clinical epidemiology. The study of outcomes of illness. New York: Oxford University Press; 1996.
31. Whittle J, Fine MJ, Joyce DZ, et al. Community-acquired pneumonia: can it be defined with claims data? Am J Med Qual 1997;12:187-93.
32. Marrie TJ, Durant H, Sealy E. Pneumonia—the quality of medical records data. Med Care 1987;25:20-4.
33. Platt R, Yokoe DS, Sands KE. Automated methods for surveillance of surgical site infections. Emerg Infect Dis 2001;7:212-6.
34. Yokoe DS, Noskin GA, Cunnigham SM, Zuccotti G, Plaskett T, Fraser VJ, et al. Enhanced identification of postoperative infections among inpatients. Emerg Infect Dis 2004;10:1924-30.
35. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. Biometrics 1983;39:207-15.
36. Greenes RA, Begg CB. Assessment of diagnostic technologies: methodology for unbiased estimation from samples of selectively verified patients. Invest Radiol 1985;20:751-6.
37. Emori TG, Edwards JR, Culver DH, Sartor C, Stroud LA, Gaunt EE, et al. Accuracy of reporting nosocomial infections in intensive-care-unit patients to the National Nosocomial Infections Surveillance System: a pilot study. Infect Control Hosp Epidemiol 1998;19:308-16.
38. Pennsylvania Health Care Cost Containment Council (PHC4). Hospital-acquired infections in Pennsylvania. PHC4 Research Briefs 2005;5:1-4.
39. Brennan PJ. In the beginning there was... heat. Infect Control Hosp Epidemiol 2006;27:329-31.
40. McKibben L, Fowler G, Horan T, Brennan PJ. Ensuring rational public reporting systems for health care-associated infections: systematic literature review and evaluation recommendations. Am J Infect Control 2006;34:142-9.